

0420
#-30-01.

#2 OIPE

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/823,676
DATE: 04/18/2001
TIME: 07:58:18

Input Set : A:\MBI0034 sequence listing.txt
Output Set: N:\CRF3\04182001\I823676.raw

ENTERED

3 <110> APPLICANT: Jiang, Cai-Zhong
5 <120> TITLE OF INVENTION: Method for Modifying Plant Biomass
7 <130> FILE REFERENCE: MBI-0034
C--> 9 <140> CURRENT APPLICATION NUMBER: US/09/823,676
C--> 9 <141> CURRENT FILING DATE: 2001-03-30
9 <160> NUMBER OF SEQ ID NOS: 8
11 <170> SOFTWARE: PatentIn version 3.0
13 <210> SEQ ID NO: 1
14 <211> LENGTH: 974
15 <212> TYPE: DNA
16 <213> ORGANISM: Arabidopsis thaliana
18 <220> FEATURE:
19 <221> NAME/KEY: CDS
20 <222> LOCATION: (62)..(874)
21 <223> OTHER INFORMATION: G1073
24 <400> SEQUENCE: 1
25 ccccccgcacc tgcctctaca gagacctgaa gattccagaa cccacactga tcaaaaataa 60
27 c atg gaa ctt aac aga tct gaa gca gac gaa gca aag gcc gag acc act 109
28 Met Glu Leu Asn Arg Ser Glu Ala Asp Glu Ala Lys Ala Glu Thr Thr
29 1 5 10 15
31 ccc acc ggt gga gcc acc agc tca gcc aca gcc tct ggc tct tcc tcc 157
32 Pro Thr Gly Gly Ala Thr Ser Ser Ala Thr Ala Ser Gly Ser Ser Ser
33 20 25 30
35 gga cgt cgt cca cgt ggt cgt cct gca ggt tcc aaa aac aaa ccc aaa 205
36 Gly Arg Arg Pro Arg Gly Arg Pro Ala Gly Ser Lys Asn Lys Pro Lys
37 35 40 45
39 cct ccg acg att ata act aga gat agt cct aac gtc ctt aga tca cac 253
40 Pro Pro Thr Ile Ile Thr Arg Asp Ser Pro Asn Val Leu Arg Ser His
41 50 55 60
43 gtt ctt gaa gtc acc tcc ggt tgc gac ata tcc gag gca gtc tcc acc 301
44 Val Leu Glu Val Thr Ser Gly Ser Asp Ile Ser Glu Ala Val Ser Thr
45 65 70 75 80
47 tac gcc act cgt cgc gcc tgc ggc gtt tgc att ata agc ggc acg ggt 349
48 Tyr Ala Thr Arg Arg Gly Cys Gly Val Cys Ile Ile Ser Gly Thr Gly
49 85 90 95
51 gcg gtc act aac gtc acg ata cgg caa cct gcg gct ccg gct ggt gga 397
52 Ala Val Thr Asn Val Thr Ile Arg Gln Pro Ala Ala Pro Ala Gly Gly
53 100 105 110
55 ggt gtg att acc ctg cat ggt cgg ttt gac att ttg tct ttg acc ggt 445
56 Gly Val Ile Thr Leu His Gly Arg Phe Asp Ile Leu Ser Leu Thr Gly
57 115 120 125
59 act gcg ctt cca ccg cct gca cca ccg gga gca gga ggt ttg acg gtg 493
60 Thr Ala Leu Pro Pro Pro Ala Pro Pro Gly Ala Gly Gly Leu Thr Val
61 130 135 140
63 tat cta gcc gga ggt caa gga caa gtt gta gga ggg aat gtg gct ggt 541
64 Tyr Leu Ala Gly Gly Gln Gly Gln Val Val Gly Gly Asn Val Ala Gly
65 145 150 155 160

RAW SEQUENCE LISTING

DATE: 04/18/2001

PATENT APPLICATION: US/09/823,676

TIME: 07:58:18

Input Set : A:\MBI0034 sequence listing.txt

Output Set: N:\CRF3\04182001\I823676.raw

```

67 tcg tta att gct tcg gga ccg gta gtg ttg atg gct gct tct ttt gca      589
68 Ser Leu Ile Ala Ser Gly Pro Val Val Leu Met Ala Ala Ser Phe Ala
69              165              170              175
71 aac gca gtt tat gat agg tta ccg att gaa gag gaa gaa acc cca ccg      637
72 Asn Ala Val Tyr Asp Arg Leu Pro Ile Glu Glu Glu Glu Thr Pro Pro
73              180              185              190
75 ccg aga acc acc ggg gtg cag cag cag cag ccg gag gcg tct cag tcg      685
76 Pro Arg Thr Thr Gly Val Gln Gln Gln Gln Pro Glu Ala Ser Gln Ser
77              195              200              205
79 tcg gag gtt acg ggg agt ggg gcc cag gcg tgt gag tca aac ctc caa      733
80 Ser Glu Val Thr Gly Ser Gly Ala Gln Ala Cys Glu Ser Asn Leu Gln
81              210              215              220
83 ggt gga aat ggt gga gga ggt gtt gct ttc tac aat ctt gga atg aat      781
84 Gly Gly Asn Gly Gly Gly Val Ala Phe Tyr Asn Leu Gly Met Asn
85 225              230              235              240
87 atg aac aat ttt caa ttc tcc ggg gga gat att tac ggt atg agc ggc      829
88 Met Asn Asn Phe Gln Phe Ser Gly Gly Asp Ile Tyr Gly Met Ser Gly
89              245              250              255
91 ggt agc gga gga ggt ggt ggc ggt gcg act aga ccc gcg ttt tag      874
92 Gly Ser Gly Gly Gly Gly Gly Gly Ala Thr Arg Pro Ala Phe
93              260              265              270
95 agtttttagcg ttttggtgac accttttgggt gcgtttgcgt gtttgacctc aaactactag      934
97 gctactagct atagcggttg cgaaatgcga atattaggtt      974
100 <210> SEQ ID NO: 2
101 <211> LENGTH: 270
102 <212> TYPE: PRT
103 <213> ORGANISM: Arabidopsis thaliana
105 <400> SEQUENCE: 2
107 Met Glu Leu Asn Arg Ser Glu Ala Asp Glu Ala Lys Ala Glu Thr Thr
108 1              5              10              15
111 Pro Thr Gly Gly Ala Thr Ser Ser Ala Thr Ala Ser Gly Ser Ser Ser
112              20              25              30
115 Gly Arg Arg Pro Arg Gly Arg Pro Ala Gly Ser Lys Asn Lys Pro Lys
116              35              40              45
119 Pro Pro Thr Ile Ile Thr Arg Asp Ser Pro Asn Val Leu Arg Ser His
120              50              55              60
123 Val Leu Glu Val Thr Ser Gly Ser Asp Ile Ser Glu Ala Val Ser Thr
124 65              70              75              80
127 Tyr Ala Thr Arg Arg Gly Cys Gly Val Cys Ile Ile Ser Gly Thr Gly
128              85              90              95
131 Ala Val Thr Asn Val Thr Ile Arg Gln Pro Ala Ala Pro Ala Gly Gly
132              100              105              110
135 Gly Val Ile Thr Leu His Gly Arg Phe Asp Ile Leu Ser Leu Thr Gly
136              115              120              125
139 Thr Ala Leu Pro Pro Pro Ala Pro Pro Gly Ala Gly Gly Leu Thr Val
140              130              135              140
143 Tyr Leu Ala Gly Gly Gln Gly Gln Val Val Gly Gly Asn Val Ala Gly
144 145              150              155              160
147 Ser Leu Ile Ala Ser Gly Pro Val Val Leu Met Ala Ala Ser Phe Ala

```

RAW SEQUENCE LISTING

DATE: 04/18/2001

PATENT APPLICATION: US/09/823,676

TIME: 07:58:18

Input Set : A:\MBI0034 sequence listing.txt

Output Set: N:\CRF3\04182001\I823676.raw

```

148                               165                               170                               175
151 Asn Ala Val Tyr Asp Arg Leu Pro Ile Glu Glu Glu Glu Thr Pro Pro
152                               180                               185                               190
155 Pro Arg Thr Thr Gly Val Gln Gln Gln Gln Pro Glu Ala Ser Gln Ser
156                               195                               200                               205
159 Ser Glu Val Thr Gly Ser Gly Ala Gln Ala Cys Glu Ser Asn Leu Gln
160                               210                               215                               220
163 Gly Gly Asn Gly Gly Gly Gly Val Ala Phe Tyr Asn Leu Gly Met Asn
164 225                               230                               235                               240
167 Met Asn Asn Phe Gln Phe Ser Gly Gly Asp Ile Tyr Gly Met Ser Gly
168                               245                               250                               255
171 Gly Ser Gly Gly Gly Gly Gly Ala Thr Arg Pro Ala Phe
172                               260                               265                               270
175 <210> SEQ ID NO: 3
176 <211> LENGTH: 1040
177 <212> TYPE: DNA
178 <213> ORGANISM: Arabidopsis thaliana
180 <220> FEATURE:
181 <221> NAME/KEY: CDS
182 <222> LOCATION: (82)..(879)
183 <223> OTHER INFORMATION: G2789
186 <400> SEQUENCE: 3
187 ctttagggac accaaatcta ttcaacctaa aagccttctt ttcccctata ttgaccaact 60
189 ttttagcgaa tcagaagagg a atg gat gag gta tct cgt tct cat aca ccg 111
190                               Met Asp Glu Val Ser Arg Ser His Thr Pro
191                               1                               5                               10
193 caa ttt cta tca agt gat cat cag cac tat cac cat caa aac gct gga 159
194 Gln Phe Leu Ser Ser Asp His Gln His Tyr His His Gln Asn Ala Gly
195                               15                               20                               25
197 cga caa aaa cgc ggc aga gaa gaa gaa gga gtt gaa ccc aac aat ata 207
198 Arg Gln Lys Arg Gly Arg Glu Glu Glu Gly Val Glu Pro Asn Asn Ile
199                               30                               35                               40
201 ggg gaa gac cta gcc acc ttt cct tcc gga gaa gag aat atc aag aag 255
202 Gly Glu Asp Leu Ala Thr Phe Pro Ser Gly Glu Glu Asn Ile Lys Lys
203                               45                               50                               55
205 aga agg cca cgt ggc aga cct gct ggt tcc aag aac aaa ccc aaa gca 303
206 Arg Arg Pro Arg Gly Arg Pro Ala Gly Ser Lys Asn Lys Pro Lys Ala
207                               60                               65                               70
209 cca atc ata gtc act cgc gac tcc gcg aac gcc ttc aga tgt cac gtc 351
210 Pro Ile Ile Val Thr Arg Asp Ser Ala Asn Ala Phe Arg Cys His Val
211 75                               80                               85                               90
213 atg gag ata acc aac gcc tgc gat gta atg gaa agc cta gcc gtc ttc 399
214 Met Glu Ile Thr Asn Ala Cys Asp Val Met Glu Ser Leu Ala Val Phe
215                               95                               100                               105
217 gct aga cgc cgt cag cgt ggc gtt tgc gtc ttg acc gga aac ggg gcc 447
218 Ala Arg Arg Arg Gln Arg Gly Val Cys Val Leu Thr Gly Asn Gly Ala
219                               110                               115                               120
221 gtt aca aac gtc acc gtt aga caa cct ggc gga ggc gtc gtc agt tta 495
222 Val Thr Asn Val Thr Val Arg Gln Pro Gly Gly Gly Val Val Ser Leu

```

RAW SEQUENCE LISTING

DATE: 04/18/2001

PATENT APPLICATION: US/09/823,676

TIME: 07:58:18

Input Set : A:\MBI0034 sequence listing.txt

Output Set: N:\CRF3\04182001\I823676.raw

```

223      125      130      135
225 cac gga cgg ttt gag att ctt tct ctc tcg ggt tcg ttt ctt cct cca      543
226 His Gly Arg Phe Glu Ile Leu Ser Leu Ser Gly Ser Phe Leu Pro Pro
227      140      145      150
229 ccg gca cca cca gct gcg tct ggt tta aag gtt tac tta gcc ggt ggt      591
230 Pro Ala Pro Pro Ala Ala Ser Gly Leu Lys Val Tyr Leu Ala Gly Gly
231 155      160      165      170
233 caa ggt caa gtg atc gga ggc agt gtg gtg gga ccg ctt acg gca tca      639
234 Gln Gly Gln Val Ile Gly Gly Ser Val Val Gly Pro Leu Thr Ala Ser
235      175      180      185
237 agt ccg gtg gtc gtt atg gca gct tca ttt gga aac gca tct tac gag      687
238 Ser Pro Val Val Met Ala Ala Ser Phe Gly Asn Ala Ser Tyr Glu
239      190      195      200
241 agg ctg cca cta gag gag gag gag gaa act gaa aga gaa ata gat gga      735
242 Arg Leu Pro Leu Glu Glu Glu Glu Glu Thr Glu Arg Glu Ile Asp Gly
243      205      210      215
245 aac gcg gct agg gcg att gga acg caa acg cag aaa cag tta atg caa      783
246 Asn Ala Ala Arg Ala Ile Gly Thr Gln Thr Gln Lys Gln Leu Met Gln
247      220      225      230
249 gat gcg aca tcg ttt att ggg tcg ccg tcg aat tta att aac tct gtt      831
250 Asp Ala Thr Ser Phe Ile Gly Ser Pro Ser Asn Leu Ile Asn Ser Val
251 235      240      245      250
253 tcg ttg cca ggt gaa gct tat tgg gga acg caa cga ccg tct ttc taa      879
254 Ser Leu Pro Gly Glu Ala Tyr Trp Gly Thr Gln Arg Pro Ser Phe
255      255      260      265
257 gataatatca ttgataatat aagtttgcgc ttcttattct ttttcacttt ttaccttttt      939
259 cacttttctta ggttttggtt taacggttga ttaatacctg aaggtttttg gaaaattttc      999
261 gatcggataa aaggatttat gttgcgagcc gaaacgcggc c      1040
264 <210> SEQ ID NO: 4
265 <211> LENGTH: 265
266 <212> TYPE: PRT
267 <213> ORGANISM: Arabidopsis thaliana
269 <400> SEQUENCE: 4
271 Met Asp Glu Val Ser Arg Ser His Thr Pro Gln Phe Leu Ser Ser Asp
272 1      5      10      15
275 His Gln His Tyr His His Gln Asn Ala Gly Arg Gln Lys Arg Gly Arg
276      20      25      30
279 Glu Glu Glu Gly Val Glu Pro Asn Asn Ile Gly Glu Asp Leu Ala Thr
280      35      40      45
283 Phe Pro Ser Gly Glu Glu Asn Ile Lys Lys Arg Arg Pro Arg Gly Arg
284      50      55      60
287 Pro Ala Gly Ser Lys Asn Lys Pro Lys Ala Pro Ile Ile Val Thr Arg
288 65      70      75      80
291 Asp Ser Ala Asn Ala Phe Arg Cys His Val Met Glu Ile Thr Asn Ala
292      85      90      95
295 Cys Asp Val Met Glu Ser Leu Ala Val Phe Ala Arg Arg Arg Gln Arg
296      100      105      110
299 Gly Val Cys Val Leu Thr Gly Asn Gly Ala Val Thr Asn Val Thr Val
300      115      120      125

```

RAW SEQUENCE LISTING

DATE: 04/18/2001

PATENT APPLICATION: US/09/823,676

TIME: 07:58:18

Input Set : A:\MBI0034 sequence listing.txt

Output Set: N:\CRF3\04182001\I823676.raw

```

303 Arg Gln Pro Gly Gly Gly Val Val Ser Leu His Gly Arg Phe Glu Ile
304      130      135      140
307 Leu Ser Leu Ser Gly Ser Phe Leu Pro Pro Pro Ala Pro Pro Ala Ala
308 145      150      155      160
311 Ser Gly Leu Lys Val Tyr Leu Ala Gly Gly Gln Gly Gln Val Ile Gly
312      165      170      175
315 Gly Ser Val Val Gly Pro Leu Thr Ala Ser Ser Pro Val Val Val Met
316      180      185      190
319 Ala Ala Ser Phe Gly Asn Ala Ser Tyr Glu Arg Leu Pro Leu Glu Glu
320      195      200      205
323 Glu Glu Glu Thr Glu Arg Glu Ile Asp Gly Asn Ala Ala Arg Ala Ile
324      210      215      220
327 Gly Thr Gln Thr Gln Lys Gln Leu Met Gln Asp Ala Thr Ser Phe Ile
328 225      230      235      240
331 Gly Ser Pro Ser Asn Leu Ile Asn Ser Val Ser Leu Pro Gly Glu Ala
332      245      250      255
335 Tyr Trp Gly Thr Gln Arg Pro Ser Phe
336      260      265
339 <210> SEQ ID NO: 5
340 <211> LENGTH: 1130
341 <212> TYPE: DNA
342 <213> ORGANISM: Arabidopsis thaliana
344 <220> FEATURE:
345 <221> NAME/KEY: CDS
346 <222> LOCATION: (189)..(1019)
347 <223> OTHER INFORMATION: G1945
350 <400> SEQUENCE: 5
351 atttcccaaa gggatttacg aaaagtcctt ctctctatc atctctttat tcaccccata      60
353 ccaacaacct ctacatcttc ttcttcttct tctctctctt ttattttctt tttaaatcat      120
355 ttacacaaaa atccaaagac aaatctgaaa tctctaataa acaaatccat aaaataagaa      180
357 aaacaaag atg aaa ggt gaa tac aga gag caa aag agt aac gaa atg ttt      230
358      Met Lys Gly Glu Tyr Arg Glu Gln Lys Ser Asn Glu Met Phe
359      1      5      10
361 tcc aag ctt cct cat cat caa caa caa cag caa caa caa caa caa caa      278
362 Ser Lys Leu Pro His His Gln Gln Gln Gln Gln Gln Gln Gln Gln Gln
363 15      20      25      30
365 cac tct ctt acc tct cac ttc cac ctc tcc tcc acc gta acc ccc acc      326
366 His Ser Leu Thr Ser His Phe His Leu Ser Ser Thr Val Thr Pro Thr
367      35      40      45
369 gtc gat gac tcc tcc atc gaa gtg gtc cga cgt cca cgt ggc aga cca      374
370 Val Asp Asp Ser Ser Ile Glu Val Val Arg Arg Pro Arg Gly Arg Pro
371      50      55      60
373 cca ggt tcc aaa aac aaa cct aaa cca ccc gtc ttc gtc aca cgt gac      422
374 Pro Gly Ser Lys Asn Lys Pro Lys Pro Pro Val Phe Val Thr Arg Asp
375      65      70      75
377 acc gac cct cct atg agt cct tac atc ctc gaa gtt cct tca gga aac      470
378 Thr Asp Pro Pro Met Ser Pro Tyr Ile Leu Glu Val Pro Ser Gly Asn
379      80      85      90
381 gac gtc gtc gaa gcc atc aac cgt ttc tgc cgc cgt aaa tcc atc gga      518

```

VERIFICATION SUMMARY

DATE: 04/18/2001

PATENT APPLICATION: US/09/823,676

TIME: 07:58:19

Input Set : A:\MBI0034 sequence listing.txt

Output Set: N:\CRF3\04182001\I823676.raw

L:9 M:270 C: Current Application Number differs, Replaced Current Application No

L:9 M:271 C: Current Filing Date differs, Replaced Current Filing Date